Using Conda Environments for Machine Learning

UPDATE: The conda environments are now supported under an Anaconda site license. In this article, the command line for activating an environment using c-shell (csh) has been updated to support this change.
We provide multiple conda environments for you to use for your machine learning projects. The environments are available via the conda module in the **`/swbuild/analytix/tools/modulefiles`** directory.

All of the available environments include the basic machine learning packages, as well as common image processing and natural language processing packages. You can activate the environments via an interactive session or with a PBS batch job.

For a list of environments we currently provide, along with the packages they include, see Machine Learning: Overview.

Loading the Module and Activating the Environments in Interactive Mode

First, access and load the miniconda module, which provides access to the environment:

```
% module use -a /swbuild/analytix/tools/modulefiles
% module load miniconda3/v4
```

Then, activate one of the environments (where ***`env_name`*** is the name of the environment):

- For bash: **`% source activate env_name`**
- For csh: **`% source /swbuild/analytix/tools/miniconda3_220407/bin/activate.csh env_name`**

Useful Conda Commands

After loading the module, you will have access to conda commands, including:

`% conda info --envs`
: Shows available environments.

`% conda list -n env_name`
: Shows installed packages within an environment.

`(env_name)% conda deactivate`
: Deactivates an environment after loading.

For more detailed documentation, see the Conda website.

Using a PBS Batch Job to Activate an Environment

You can activate your machine learning environment, run your program, and deactivate the environment in a PBS script. For example:

```
#!/bin/bash -x
#PBS -l select=1:model=sky_gpu:mpiprocs=1:ncpus=36:ngpus=4:mem=300g
#PBS -l place=scatter:excl
#PBS -q v100
#PBS -l walltime=1:00:00
#PBS -N test
```

```
#PBS -j oe

cd $PBS_O_WORKDIR

module purge

# load the module and environment
module -a use /swbuild/analytix/tools/modulefiles
module load miniconda3/v4
source activate env_name

# run python script
python test_cnn.py

# deactivate environment
conda deactivate

# end of script
```

For more detailed information about running your PBS job on the GPU nodes, see Requesting GPU Resources.

Requesting New or Missing Packages

If you need to use a specific package that is not currently installed in any of the environments, please send an email to support@nas.nasa.gov. In your email, provide the name of the package and request a ticket be opened with the Data Science team. We will help install the package into the existing environments or create a new environment if needed. You can also install your own packages or create your own conda environment.

Article ID: 557
Last updated: 14 Jun, 2022
Revision: 126
Machine Learning -> Machine Learning at NAS -> Using Conda Environments for Machine Learning
https://www.nas.nasa.gov/hecc/support/kb/entry/557/